DOCUMENT RESUME

ED 127 348 TM 005 437

AUTHOR Marshall, Hermine H.; And Others
TITLE Stability of Teacher Behaviors as Measured by a Broad Range Low-Inference Observational System.
SPONS AGENCY California Univ., Berkeley. School of Education.
PUB DATE [Apr 76]
NOTE 58p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Behavior Rating Scales; *Classroom Observation Techniques; Classroom Techniques; Effective Teaching; Open Education; Primary Education; *Reliability; Self Contained Classrooms; Statistical Analysis; *Teacher Behavior; Teaching Styles; Traditional Schools
IDENTIFIERS *Dimensional Occurrence Scale

ABSTRACT

Patterns of stability of behaviors in classrooms with varying types of structure were investigated by means of a low-inference rating scale applicable to both traditional and nontraditional classrooms. Utilizing a research design which controls for subject matter, period of day, and day of week, eighteen observations were made in each of six primary classrooms. Behavioral items were combined into variables reflecting cognitive, affective, interpersonal, and structural aspects of the classroom. Results indicate that certain classroom variables were stable, while others varied according to period of day, subject matter (activity), and type of classroom structure. Teachers differed in their patterns of stable behavior. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED127348

STABILITY OF TEACHER BEHAVIORS AS MEASURED BY A BROAD RANGE LOW-INFERENCE OBSERVATIONAL SYSTEM[1]

Hermine H. Marshall, Judith L. Green,[2]
and Maya Tsūji Lawrence
University of California, Berkeley

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Purpose

The purpose of this paper is to examine questions of (1) whether the behaviors and strategies of individual teachers vary or remain constant according to subject matter, period of day, and day of week; and (2) whether the patterns of stable behavior vary for teachers with differing classroom structures -- ranging from the traditional to the "open."[3]

This study was designed to overcome two types of problems which have been plaguing research on teacher effectiveness: (1) the lack of observational systems which are equally applicable to a range of classroom structures, extending from the traditional to the open; and (2) the lack of evidence regarding the stability of teaching behaviors and strategies.

Applicability of Observational Systems

Most observational systems used in previous research focus on the more traditional elements of the classroom, e.g. "T lectures," "St recites." Where attempts have been made to develop and validate observational systems designed for open classrooms, a variety of problems have been encountered. Some scales, e.g. the Walberg & Thomas scale (1972), have been constricted by an undue weighting of items reflecting the more surface aspects of the classroom, such as types of materials, grouping of students (see Marshall, 1975). Investigations utilizing

TM005 437

rating systems which require inferential judgments, e.g. the Classroom Atmosphere Ratings (Solomon, 1973), are usually handicapped by the halo effect common to high-inference rating scales. Where sign systems have been explored, the inability to accurately represent the frequency and intensity of the observed behaviors has hindered the differentiation of types of classrooms (Marshall, 1974). The use of verbal interaction analysis systems -- even where modified to more accurately represent the variety of verbal interaction patterns found in a range of types of classrooms -- does not account for non-verbal behaviors and global structuring strategies (Marshall & Green, 1976). Consequently, in order to measure the broad variety of behaviors found in both traditional and open types of classrooms (Bussis and Chittenden, 1970; Marshall, 1972; Spodek and Walberg, 1975), two previous observation systems -- the Classroom Atmosphere Ratings (Solomon, 1973) and the Dimension Sign System (Marshall, 1973) -- were modified and synthesized to form the Dimensional Occurrence Scale (DOS). Because the items on the DOS were derived from criteria based on a review of the Open Education literature (Marshall, 1972), many of the DOS items are similar to those on the Walberg and Thomas Scale (1972) and on the Dimensions of Schooling (DISC) questionnaire (Musella, 1973). The process of item derivation and refinement is elaborated in the Interim Reports (Marshall, December 1973; December 1974).

The DOS is used to rate the occurrence of clearly observable behaviors which require the observer to make a minimal amount of inference regarding the occurrence or interpretation of these behaviors. As suggested by Dunkin and Biddle (1975), the low-inference behavioral rating items on the DOS have been combined into higher inference variables which reflect the broader cognitive, affective, and structural aspects of the classroom.[4] Table 1 lists the variables together with the low-inference behavioral items comprising each of the variables.

Because the DOS is used to rate the occurrence of low-inference behaviors, the halo effect common to high inference ratings scales is avoided, yet the frequency and intensity of the observed behaviors can be indicated by the points on the rating scale. Furthermore, incorporating a broad range of items eliminates (1) the emphasis on surface elements of the environment -- which has limited the usefulness of the Walberg & Thomas Scales, and (2) the exclusive use of verbal behavior which restricts the applicability of interaction analysis systems.

Stability of Classroom Variables

Although a recent study (Brophy, Coulter, Crawford, Evertson, King, 1975) has found overall stability for many classroom variables, most reviews of the literature have reported a lack of evidence regarding the stability of classroom variables (Medley & Mitzel, 1963; Rosenshine, 1970). Two major factors seem to be involved in this apparent lack of stability: (1) the meaning of the term "stability" and its implications both for classroom research and for classroom practice, and (2) inadequacies in the experimental design of much of the previous research.

Meaning of Stability

As used in this paper, the term "stability" refers to a consistency in the frequency of classroom behaviors. That is, behaviors which are stable are rated at the same or a similar rating point (degree) during each observation period. It should be noted, however, that the opposite end of the "stability" continuum does <u>not</u> imply instability or inconsistency, but rather variability, fluctuation, or flexibility. From a pedagogical point of view, this flexibility usually has a deliberate purpose. Fluctuations in the frequency of particular

behaviors or strategies may be a function of changes in the classroom structure for different subject matters or for different educational purposes. For example, certain behaviors may occur in the context of a "Class meeting" or "Sharing time" that would not generally occur during a math lesson. For certain teachers, changes in the occurrence of certain behaviors may coincide with changes in the period of the day. For example, Reading may occur during the second period of the morning each day. The teacher may use certain strategies during this period which s/he does not use during the first period when there is Sharing or News time. The occurrence of certain behaviors may also change within the same time period but on different days of the week. For example, in certain classrooms, Spelling may be scheduled during the third period on Monday, Wednesday and Friday; while Art is scheduled for third period on Tuesday and Thursday. A different set of behaviors would be expected during these two activities. If period 3 on Mondays and Wednesdays is examined, the behaviors may appear quite stable. In contrast, inspection of period 3 behaviors on Mondays and Thursdays may indicate a lack of stability -- i.e. flexibility -- which reflects differences in the ongoing activities.

In some classrooms, often labelled "open," students may be engaged in a variety of different activities simultaneously during a particular period or throughout the entire day. In these classrooms, the particular activities which occur may vary somewhat from day to day. Consequently, both the teacher and the students may be observed to exhibit varying behaviors both within the same period and from day to day, depending on the particular activities in which the individuals are engaged. The apparent lack of stability of classroom behaviors may in reality reflect the range of behaviors appropriate to the needs of the individual students involved in different activities at these particular times.

On the other hand, high stability may be an indication of a limited repertoire of teaching strategies or even of rigidity. Therefore, in actual classroom

practice, high stability within periods and across periods and days may not necessarily be a characteristic of an effective classroom. A certain amount of variability is more likely to reveal behavior that adapts to the changing needs of the situation and the students.

Experimental Design

In contrast, stable variables are needed for research purposes. However, much of the prior research has failed to control for such situational differences as subject matter and time of day. The failure to control for these situational variables in previous research may have contributed to measurement error and the appearance of instability (McGaw, Wardrop, & Bunda, 1972; Frick & Semmel, 1974), which in turn may have led to the paucity of significant relationships between classroom variables and student outcomes (Dunkin & Biddle, 1974; Emmer & Peck, 1973; Rosenshine & Furst, 1973). This study was designed to control for these situational factors and to facilitate an analysis of the stability of behaviors in both traditional and nontraditional classrooms.

In planning for the data collection, an attempt was made to gather data during the Reading period and two other periods each day in each classroom. In all the classrooms except classroom K, Reading occurred during period 1; classroom K had Reading during period 2. Where other activities in addition to Reading occurred during Reading period, additional strategies would be expected to be used and thus attenuate the stability of the behaviors. For most teachers, Reading as the sole content of the learning activity occurred on several days – at least during the first part of the Reading period when the T-DOS was used. Frequently, though, other activities occurred either simultaneously with the Reading activities or after the students completed their Reading assignment. Because of the reality of the dynamic classroom situation, the subsequent

analysis of the stability of the classroom variables within the Reading period was confounded by the occurrence of other activities during this period. The effect of a variety of activities occurring during a single period is considered below in the Results section on Apparent Lack of Stability.

Research Questions

The specific research questions which this aspect of the study was designed to answer are as follows:

1. Which classroom variables are stable within and across periods?
2. Which classroom variables fluctuate within and across periods?
3. What factors contribute to the fluctuation of classroom variables?
4. Is the stability of classroom variables related to the type of structure of the classroom? For example, are variables in traditional classrooms more stable within and between periods than in more open classrooms where a greater range of activities occur simultaneously?

Method

Instrument

The Dimensional Occurrence Scale is divided into two sections: (1) the Teacher DOS (64 items) and (2) Student DOS (54 items). This format allows the observer to attend first to the teacher for a 16-minute observation period and then to the student for a 16-minute observation period, thus providing two vantage points from which to examine the teacher's behavior and purpose.[5]

The rating items require the observer to make a minimal amount of inference regarding the occurrence or interpretation of the behaviors included. The behaviors are rated on one of the three types of 5-point scales: (a) the amount

of time during the observation period that the behavior occurs, ranging from 1-"no occurrence of behavior observed" to 5-"behavior occurred during the entire period;" (b) the number of times certain infrequently occurring behaviors occurred, ranging from 1-"no occurrence observed" to 5-"seven or more occurrences of the behavior were observed;" and (c) the number of students exhibiting the behavior (St-DOS only), ranging from 1-"less than one-fourth of the students present exhibited the behavior" to 5-"three-fourths or more of the students present exhibited the behaviors." A complete description of the DOS and procedures for use are given in the Manual (Marshall, 1976a).

Interrater agreement (reliability) between three trained observers was calculated for each item using three methods: (1) Ebel's intraclass correlation (Guilford, 1954), (2) a measure of explained variance (Marascuilo, 1971), and (3) percent of agreement. Six unreliable items -- all on the T-DOS - were excluded from subsequent data analysis. The procedures for determining interrater agreement and the decision to exclude items where the interrater agreement was unsatisfactory are elaborated in the Appendix.

Subjects

Six primary classrooms (grades 1-3) from a middle class suburban area were selected by district administrators as outstanding examples of different types of classroom structures. The classroom structures ranging from "open" (defined as using "flexible learning centers and a more individualized and interdisciplinary approach") to "traditional" (defined as using "a high degree of structure as exemplified by two or three groups for reading, spelling, and math"). The teachers were also interviewed by both the Principal Investigator and an outside interviewer to obtain a picture of the teacher's perception of her classroom structure and viewpoint on teaching.

Design

Eighteen observations were made by three trained observers with the T-DOS in each of the six classrooms. Each classroom was observed on the same two days of the week for three consecutive weeks. During each of the six mornings, three 16-minute observations were made with the T-DOS at prespecified times. Twelve observations were made in each classroom with the St-DOS, two 16-minute observations during each of the six mornings. (See Table 2.) Because of the timing of recesses and lunch, it was impossible to schedule a third observation with the St-DOS.

Results and Discussion

Data Analysis Procedures

Classroom behaviors have been analyzed according to (1) the stability of the variables in terms of the relative frequency with which the items comprising each variable are observed to occur; (2) the stability of occurrence of each individual item within the variable; and (3) differences between teachers in patterns of consistency and flexibility.

Stability of the Variables

The stability of the variables listed in Table 1 has been analyzed by investigating whether the relative frequency (rank order) with which the component items occur remains constant for each teacher within and across periods. For example, within the Grouping variable, does Teacher X consistently utilize whole class grouping with greatest frequency, large group structure second most frequently, and individual attention least frequently? That is, regardless of whether each of these items always receives the exact same rating, does

the relative order of frequency among these items remain constant within and across periods?

The stability of the variables in terms of the relative frequency of occurrence of the component items was analyzed by means of Friedman's analysis of variance by ranks (Siegel, 1956). Where the relationships were significant, Kendall's Coefficient of Concordance (W) was calculated as a measure of explained variance. Where the results were nonsignificant, the raw data was inspected to determine the factors influencing the apparent lack of stability.

In the analysis of the stability of the variables which follows, first the general stability across all periods in the morning is described. Then the stability of the variables within periods is examined. And finally, reasons for the apparent lack of stability are considered.

The results of Friedman's analysis of variance by ranks and Kendall's Coefficient of Concordance (W) are displayed in Table 3 (T-DOS) and Table 4 (St-DOS).

Since there are 145 significant correlation coefficients out of a possible 264 on the T-DOS, and 143 significant correlation coefficients out of a possible 198 on the St-DOS, chance expections indicate that approximately 7 of these on each section of the DOS are Type I errors.

A statistically significant χ_r^2 is only one indication of the stability of the variables, however. Inspection of the raw data reveals that the failure of certain of the variables to reach statistical significance by Friedman's analysis of variance by ranks is often due to the lack of variability in the ratings of the items composing the variable. For example, where almost all the items within a variable were rated "1" or "1 or 2," this lack of variability in the rating results in a low and non-significant χ_r^2. These statistically non-significant results are misleading therefore, since teachers who

who are rated with almost all 1's or 1's and 2's on the component items are in reality quite stable in their use of these behaviors.

Consequently, in order to arrive at a more accurate picture of the stability with which behaviors are observed in the classroom, it is necessary to add to the statistically significant variables those variables for which there is a minimal range of rating scores. Table 5 shows the stability of the variables across all periods as indicated by both (1) Friedman's analysis of variance by ranks and (2) minimal variability of component items. Table 6 (T-DOS) and Table 7 (St-DOS) present the same information for the stability of the variables within periods.

General Stability of Variables across Periods across All Teachers

Across all periods, Friedman's Analysis of Variance by Ranks indicated that the χ_r^2's for 7 of the 11 variables on the T-DOS and 7 of the 11 variables on the St-DOS were statistically significant. See Tables 4 and 6. These variables are Grouping (T-DOS and St-DOS), Orientation towards Learning (T-DOS and St-DOS), Responsibility for Learning (T-DOS and St-DOS), Control (T-DOS and St-DOS), Acceptance (T-DOS), General Tone (T DOS and St-DOS), Orientation toward Individual (T-DOS and St-DOS), and Flexibility (St-DOS). However, inspection of the raw data reveals that for 3 variables on both the T-DOS and the St-DOS, the lack of a significant χ_r^2 can be attributed to the infrequent occurrence of the items comprising the variables. Where most ratings were 1 or 2, behaviors were in reality consistent although this stability is not indicated by the Friedman analysis. Consequently, Evaluation (T-DOS and St-DOS), Affect (T-DOS), Cooperation (St-DOS), Involvement (St-DOS), Flexibility (T-DOS), are in reality also stable. (See Table 5.)

Only for Subject/Material does the lack of statistically significant χ_r^2 appear to reflect real variability -- at least on the T-DOS. Although the χ_r^2 was statistically significant on the T-DOS for two of the teachers for the Subject/Material variable, the Coefficient of Concordance (W) was low for these teachers ($< .46$) -- suggesting that only a small proportion of the relationship is accounted for.

In other words, inspection of the data shows that the classroom structure as indicated by the items on the Subject/Material variable often does in reality change from one period to another during the morning.

Thus, based on the statistical analysis of the relative frequency with which the behaviors comprising the variable occur across all periods and on an inspection of the raw data, all of the variables except Subject/Material appear to be generally stable across all periods for all teachers. The fluctuation of the Subject/Material variable seems to reflect a change in the structure during the morning for different activities.

General Stability within Periods across Teachers

Within each period across days, the χ_r^2 reached significance ($p < .05$) for all teachers for only one variable: Responsibility for Learning on the St-DOS -- although the χ_r^2 for General Tone on the St-DOS lacked only one instance of reaching the .05 level and that one instance was significant at the .10 level. None of the χ_r^2's for the variables on the T-DOS reached statistical significance for all teachers and within each period.

However, where the χ_r^2 did not reach statistical significance, inspection of the raw data indicated that for four variables on the T-DOS and for one variable on the St-DOS, the non-significant χ_r^2's were due to minimum range of rating scores and/or infrequent occurrence of the component items. These

variables on the T-DOS are Orientation towards Learning, Control, General Tone, and Flexibility. For the St-DOS, the variable which shows minimal range of rating is Evaluation. Consequently, these variables are in reality stable within period. For the remaining variables, some teachers are consistent in the relative frequency with which they utilize the behaviors comprising these variables during some periods; while in other periods, there is greater fluctuation in the use of the component items -- as is elaborated in the next section.

Tables 6 and 7 summarize those variables which are stable within periods as evidenced by either (1) a significant χ_r^2 or (2) minimum range of rating scores.

Lack of Apparent Stability of Variables Due to Differences in Activities within Periods

In many cases, a within-period analysis of the stability of variables may be inadequate either because of (1) a planned difference in activities within the same period but on different days -- as exemplified by the Mon-Wed-Fri. Spelling, Tues-Thurs. Art type of structure, (2) a classroom structure such that multiple activities occur at the same time during one or more periods -- e.g. the learning center approach, or (3) a structure where students who finish their assignment either begin something else or become restless -- particularly towards the end of a period.

Where multiple activities occur simultaneously or where additional activities occur when students finish an assignment, a precise analysis of the apparent lack of stability is, for the most part, impossible. Where many activities occur during the same period, the apparent lack of stability in the ratings is most likely due to the differential use of behaviors appropriate to the varied activities.

However, where the activities within periods change according to the day of the week, a further analysis is possible. A descriptive analysis of the fluctuation in the ratings of the items comprising the variables due to these activity changes follows. (The small sample of observations for activities other than Reading precludes a statistical analysis.)

Those variables for which inspection of the raw data reveals that the lack of statistical significance of the χ_r^2 is due to a real difference in item rating associated with differing subject matter activities, tasks, or structure on different days even within the same period are also noted in Tables 6 and 7.

These differences in ratings due to changes in subject matter and structure of activity can be exemplified by Teacher A. (See Table 8.) During period 3, on two of six days, the activity was "Sharing," while on the remaining four days the activity was "Centers." Differences stemming from the different activities during this period are reflected on the T-DOS ratings as follows: (1) "individual attention-cognitive" (Grouping item) was not observed to occur during Sharing, while it was observed during Centers; (2) no evaluation strategies occurred during Sharing but evaluation behaviors did occur during Centers; (3) more "encouragement of open expressiveness" (Acceptance item) occurred during Sharing than during Centers; and (4) "noisy" (Control item) was observed only during Sharing.

Or for Teacher E during period 2 on three days the activity was Sharing, while on the remaining three days the activity was Writing--although on one of the latter days Language and Spelling also occured. (See Table 9.) On the days when Writing alone occurred, item #16 Same materials (Subject/Materials) received a high rating. Item #15 Receptive activities (Orientation Towards Learning) received a high rating during Sharing, but did not occur during the Writing period. On the Responsibility for Learning variable, three items occurred frequently during Writing but not during Sharing: #18 Moves freely

for learning, #87 Seeks aid of T, and #9 Works unaided. Expression of feelings (item #107 on Affect variable) occurred only during Sharing, not during Writing. On the Involvement variable, item #73: Sts uninvolved in activity, indicated that half the class was uninvolved during Sharing but not during Writing. On the composite variables Orientation toward the Individual and Flexibility, items indicating movement occur during Writing period, but only infrequently during Sharing time.

These differences in ratings which are a function of different activities within the same period but on different days result in different orders of relative frequency of the items comprising these variables. The lack of apparent stability within period is thus due to changes in activity rather than to nonstability of behaviors. Such examples illustrate that random sampling without noting the activity in the classroom--even random sampling within period--is likely to lead to a false impression of inconsistency of classroom variables.

Nevertheless, for some variables, no clear-cut reason for the lack of stability emerges from inspection of the raw data. It appears that in these cases, the teacher's strategies and students' behaviors change for reasons that the DOS and this method of data collection do not reveal.

Item Stability

The stability of occurrence of each item has also been examined for each teacher within and across periods. Stability of individual items refers to whether the rating point for each item remains constant from one observation period to the next. For example, does Teacher Y always give praise three or four times during each period? Or does the amount of praise given by Teacher Y vary widely from one observation to the next?

The stability of each item within and across periods, within and across

teachers was determined by compiling into a table the number of times each item was rated at each level and comparing the results within and across periods, within and across teachers.

Table 10 (T-DOS) and Table 11 (St-DOS) show the range of ratings for each individual item comprising the variables within each period for each teacher. Items were considered stable when they were coded at the same point for five or six of the six days. Items coded with a single numeral, e.g. 2, mean that the ratings for that period were the same on each of the six days and are consequently stable items. Items with a single numeral and superscript, e.g. 2^4, indicate that the items were rated at the same point, e.g. 2, for five of the six days and at the point indicated by the superscript, e.g. 4, on the sixth day. These items, too, are generally stable--particularly where the superscript numeral varies from the main rating by only one point.

Items with a superscript in parentheses,--e.g. $1^{(4)}$ or $1^{(2,3)}$ mean that the behavior was rated at the same point for four days and received the rating in parentheses on the other two days. Items coded with a slash between two numerals, e.g. 3/5, indicate that the item was rated at one point e.g. 3, on three days, and at the other point, e.g. 5, on the remaining three days. Items coded in this manner often reflect a classroom structure which changes within period depending on the day of the week, e.g. on Monday, Wednesday, and Friday during this period, there is Sharing; while on Tuesday and Thursday, Spelling occurs. Where the ratings are more variable, the range of variability is indicated by a hyphen between the two numerals at the ends of the range. For example, 3-5 means that during that period, this item received ratings of 3,4,5. An item was considered moderately stable if it was rated within a range of three points (e.g. 1-3) or rated at the same point for four of the six days in each period for all teachers (e.g. $2^{(3,5)}$).

An item rated with a single numeral rating, e.g. 4, during period 1 and a range, e.g. 1-5, during period 2 indicated that that item was stable during period 1 but fluctuated during period 2.

In addition to being able to ascertain which items fluctuate within period and between periods for each teacher, these tables can also indicate between-teacher differences in stability as well as differences between teachers in absolute ratings, reflective of teacher style.

Although many of the classroom variables are stable in terms of the relative frequency with which the component items are utilized as indicated by the χ_r^2 and an inspection of the raw data, examination of the absolute (as contrasted with the relative) frequency of each individual item reveals few individual items which are stable within and across periods for all teachers. According to the above criteria, twelve of the 58 items on the T-DOS and seven of the 54 items on the St-DOS are stable for all teachers. Sixteen items on the T-DOS and 16 on the St-DOS were judged as moderately stable.(See Tables 10 and 11.)

With two exceptions, all of the items which are stable or moderately stable across all teachers are items which rarely or infrequently occurred. In other words, part of the reason for the stability of these items is that on most occasions, the behaviors were not observed to occur and therefore received a rating of 1. When items judged "stable" and "moderately stable" did occur they were usually rated 2 or 3.

The first exception to the low frequency of the stable items is that for teacher C--the most open of the teachers as judged by the observers and an outside interviewer--one item (#72 Sts make choice) was rated at the highest point (5) during one period, while this item was usually rated at the lowest or nonoccurrence point (1) for the other teachers. The second exception is that for one item (#67 Sts followed posted or stated sequence), all of the teachers,

except teacher C, were usually rated at the highest (5) point.

In brief, those items which can be considered stable or moderately stable across all teachers are infrequently occurring items and derive their stability in part from their infrequency or lack of occurrence. In contrast, with the exception of item #67, those behaviors which occur with more than minimal frequency occurred with a greater range of variability.

Where variability of individual items occurs, it can usually be attributed to differences in activity either within or between periods, as noted in the section on Lack of Apparent Stability of the Variables (page 12). The variability of individual items due to differences in activity was noted for teachers A and E in the analysis of the lack of stability of the relative frequency of the items comprising the variables (See page 13 and Tables 8 and 9).

Teacher K can serve as an additional example of the fluctuation of items due to changes in the activity within and between periods. Sharing was scheduled during period 1 each day; Reading occurred each day during period 2; and during period 3 on three days Discussion and Spelling were scheduled; while on the remaining three days, Writing, Math, and Individual Work occurred. During period 1 (Sharing), teacher K utilized total class structure (item #11) and common pace for class (#19) on each of the six days, while these items were variably rated during periods 2 and 3. During period 1, Individual attention (item #7) was stable in that it did not occur (rated 1). During period 2, this item was stable, but was rated at the second point on the rating scale. During period 3, total class and common pace for class were rated high only on the days when there was Discussion and Spelling. On the days when there was Writing, Math, and Individual Work, individual attention was given. (See Tables 10 and 11.)

Between Teacher Differences in Patterns of Stability

Differences between teachers in their patterns of stability of behaviors have been noted in previous sections. Between teacher differences in the stability of classroom behaviors reflecting open and traditional classroom structure can be elaborated further with regard to both (1) the stability of the variables in terms of the relative frequency with which the component items are used and (2) the stability of the individual items.

Between Teacher Difference in Stability of Variables. As can be seen in Tables 5 and 6, the variables which are stable within and across periods seem to vary somewhat between teachers, suggesting that teachers differ in their patterns of consistent behavior. For example, for most teachers, the relative frequency with which behaviors within the Evaluation variable on the T-DOS are used is quite stable. However, for Teachers A and K, there is some variability due to activity during period 3 but not during periods 1 and 2. For Teacher C, the use of these behaviors fluctuates in all periods. It may be that Teacher C is more flexible in her use of evaluation strategies than are the other teachers.

In comparing the most "open" teacher, Teacher C, with the most traditional teacher, Teacher K, the most striking differences on the T-DOS are the differences in the stability of the Grouping variable and the Subject/Material variable. Teacher K is stable in the relative frequency with which she utilized the behaviors comprising the Grouping variable in all three periods, while the grouping strategies used by Teacher C seem to be more flexible and seem to be a function of the activities during all three periods. The relative frequency of the items on the Subject/Materials variable is stable for Teacher K in two of the three periods but varies for Teacher C in all three periods. These differences between teachers are consistent with the theoretical expectations for the two types of classrooms.

On the St-DOS, differences between these two teachers are apparent on the Subject/Material, General Tone, and Involvement variables. During period 1 in classroom K, the items within the Subject/Materials variable seem to change according to the activity (Sharing and Discussion vs. Math, Science, Spelling); while they are stable during period 2 (Reading.) In classroom C, though, they are stable during both periods. During period 1 in classroom K, the General Tone and Involvement variables seems to vary according to the activity. These variables are stable during period 2 (Reading) for classroom K and during both periods for classroom C. When there was Sharing and Discussion during period 1 in classroom K, students smiled more and were more uninvolved and restless than during Math, Science and Spelling.

Between Teacher Differences in Stability of Items. Inspection of Tables 10 and 11 shows that some items are stable within or across periods for certain teachers but not for others. Other items are stable during one period but vary during other periods for certain teachers. Additional between-teacher differences are found in the degree and the range of rating points for various items.

In addition to the differences between teachers noted in the preceding section, perhaps the most apparent differences in item stability occur within the Responsibility for Learning variable on the St-DOS. See Table 11. Item #72 Sts make choice is stable at the low point (1) for all teachers except teacher C, --as noted above. Teacher C received a high and stable rating (5) for four of the six days. A similar pattern is seen for item #104 Sts make choice (although this item was rated more variably in classroom H during period 2.) And the low and high points were reversed between teacher C and the other teachers as stable points for item #67, as noted previously, Item #90 Sts plan sequence is stable at the low point (1) for all teachers except teacher C, whose rating is (1-5). These differences in patterns of stability and level of rating seem

to be attributable to the greater encouragement of students assuming responsibility for their own learning associated with "open" classroom philosophy.

Other differences reflective of differences in types of classroom structure have been analyzed by means of correlating the variables with an independent interviewer's rank order of the teachers on dimensions of Openness (Marshall, 1976b).

Summary and Conclusions

This study seems to show that when stability is considered in terms of the relative frequency of behaviors comprising the variables, there appears to be a general stability across all periods for all teachers; whereas within periods, there is greater fluctuation. When the stability of the individual behaviors is examined, the more frequently occurring behaviors show a greater amount of fluctuation compared to the stability of infrequently occurring behaviors. According to both types of analysis, it appears that variability is usually associated with (1) changes in activity from one observation period to the next and/or (2) differences in type of classroom structure.

More specifically, all of the classrooms show general stability over all of the morning's observation periods for all variables except Subject/Materials. The Subject/Material variable--as measured by the T-DOS--seems to fluctuate in most classrooms with changes in the activity from one period to the next.

Within a given period, however, most of the variables measured show greater fluctuation from one day to the next. While several variables are stable within each period for all teachers (Orientation towards Learning, Control, General Tone, and Flexibility--as measured by the T-DOS--and Responsibility for Learning, General Tone, and Evaluation--on the St-DOS), other variables seem to fluctuate within the same period from one day to the next--at least for some of the teachers. In many cases, changes in the relative frequency of occurrence of the behaviors

comprising the variable coincide with differences in activity within the same period but on different days. In addition, some of the differences between teachers in the stability of classroom variables within periods also seems to reflect the type of classroom structure the teacher has established.

When the stability of the individual behaviors comprising the variables are examined in terms of occurrence of the total frequency rather than the relative frequency within each period, it appears that it is mainly the infrequently occurring behaviors which are stable for all teachers. Most of the more frequently occurring behaviors fluctuate from one observation period to the next, at least for some teachers. This variability in the amount of occurrence of individual items similarly seems to be a function of either (1) differences in the activity during the observation period or (2) differences in the type of classroom structure.

Those behaviors which seem to vary according to the type of classroom structure--particularly in the most open classroom--include those associated with (1) the type of grouping for instruction, (2) the use of a variety of subjects and materials at the same time, and (3) the encouragement of the student's beginning to assume responsibility for his/her own learning.

Although the results of this study should be considered cautiously due to the small number of teachers in the sample, the fact that the variability of classroom behaviors appears to be associated with differences in subject matter and activity as well as with the type of classroom structure has implications for research on teacher effectiveness. Utilization of an experimental design which adequately controls and samples the type of classroom structure and the activity during the observation period will be more likely to result in the precision and stability of variables generally needed to achieve statistically significant results.

Furthermore, the differences found in patterns of stability according to the type of classroom structure suggest that the Dimensional Occurrence Scale has

potential for filling the need for an observational system applicable to both traditional and open classrooms. (Convergent and predictive validity of the DOS are currently being analyzed.)

References

Brophy J.E., Coulter, C.L., Crawford, W.J., Evertson, C.M., & King, C.E. Classroom observation scales: stability across time and context and relationships with student learning gains. Journal of Educational Psycholcgy, 1975, 67, 873-881.

Bussis, A. & Chittenden, E. Analysis of an approach to open education. Princeton, N.J.: Educational Testing Service, 1970.

Dunkin, M.J. & Biddle, B.J. The study of teaching, New York: Holt, Rhinehart, & Winston, 1974.

Emmer, E.T. & Peck, R.F. Dimensions of classroom behavior. Journal of Educational Psychology, 1973, 64, 223-240.

Frick, T. & Semmel, M.I. Observational records: observer agreement and reliabilities, Paper presented at Meeting of American Educational Research Association, Chicago, 1974.

Guilford, J.P. Psychometric methods, (2nd ed.). New York: McGraw Hill, 1954.

Marascuilo, L. Statistical methods for behavioral science research, New York: McGraw Hill, 1971.

Marshall, H.H. Criteria for an open classroom. Young Children, 1972, 28, 13-20.

Marshall, H.H. Dimensions of classroom structure and functioning project. Interim Report. (mimeo), 1973.

Marshall, H.H. Dimensions of classroom structure and functioning project. Interim Report. (mimeo), 1974.

Marshall, H.H. Clarification of open education: an analysis of the Walberg & Thomas scales. Research In Education, 1975

Marshall, H.H. Dimensional occurrence scale: manual. 1976a, (in preparation)

Marshall, H.H. Dimensions of classroom structure and functioning project. Final Report. (mimeo), 1976b

Marshall, H.H. & Green, J.L. Classroom verbal behavior: contextual purpose and situational factors. Paper presented at Meeting of American Educational Research Association, San Francisco, April, 1976.

McGaw, B., Wardrop, J.L., & Bunda, M.A. Classroom observation schemes: where are the errors? American Educational Research Journal, 1972, 9, 13-27.

Medley, D.M., & Mitzel, H.E. Measuring classroom behavior by systematic observation. In Gage, N.L. (Ed.). Handbook of Research on Teaching, Chicago: Rand McNally, 1963, 247-328.

Musella, D. Closure on openness in education: dimensions of school. Paper presented at Meeting of American Educational Research Association, New Orleons, February, 1973.

Rosenshine, B. Evaluation of classroom instruction. Review of Education Research, 1970, 40, 279-300.

Rosenshine, B. & Furst, N. The use of direct observation to study teaching. In Travers, R. (Ed.), Second Handbook of Research on Teaching, Chicago: Rand McNally, 1973.

Siegel, S. Nonparametric methods for the behavioral sciences. New York: McGraw Hill, 1956.

Solomon, D. Progress report: individual characteristics and children's performance in varied educational settings. (mimeo), n.d. (1973).

Spodek, B. & Walberg, H. Studies in open education. N.Y.: Agathon, 1975

Walberg, H., & Thomas, S. Open education: a classroom validation in Great Britain and the United States. American Educational Research Journal,1972, 9, 197-207.

Footnotes

1. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April, 1976.

The authors would like to thank Thomas A. Cesa for his help with the statistical analysis and his reading of an earlier version of this paper; and Ruth Vogt and Nancy Roark for their help in the data collection and in making suggestions for the refinement of the instrument.

2. Now at Kent State University, Ohio.

3. The study reported in this paper is a part of a multiple measure research project designed to explore the effectiveness of teacher behaviors in a variety of types of classrooms--ranging from the traditional to the "open". The Project was funded by the Instructional Laboratories, School of Education, University of California, Berkeley.

4. While the traditional way of grouping items to form variables has been to utilize a statistical procedure such as factor analysis, these techniques require large samples and are therefore inappropriate for the purposes of the present study. Hence, in order to present a picture of the surface structure of the classroom in a descriptive and understandable manner, the items on the DOS were grouped on a logical basis. An analysis of these logical groupings by means of inter-item correlations is elaborated in Marshall (1976b).

5. Note that items such as "T works with small groups' (T-DOS) and "Sts work at own pace" (St-DOS) both provide indications of the purpose and the structure that the teacher has established.

Table 1

DIMENSIONAL OCCURRENCE SCALE VARIABLES WITH COMPONENT ITEMS

Categories of Variables[1]	Variables	Items from T-DOS	Items from St-DOS
SUPPORT STRUCTURE FOR LEARNING	GROUPING	7. T gives individual attention for cognitive purposes 22. T uses flexible grouping for instruction 24. T works with small groups 14. T works with large groups 17. T presents information to group 5. T uses pre-set grouped instruction 11. Total class activities 19. Common pace for whole class 23. Common pace for groups 31. Rushed	85. Own pace (#) 77. Individualized instruction 92. Individualized instruction (#)[2] 65. Spontaneous grouping 102. Spontaneous grouping (#)
	STRUCTURE OF SUBJECT MATTER & MATERIALS	28. Varied subjects 29. Varied materials 10. T works with Sts using text/workbook approach to skill and concept learning 6. Teacher used supplementary materials	117. Sts used diverse materials 118. Varied subjects 116. All Sts worked with same materials, books at same time 99. Sts use text/workbook approach (#) 93. Sts use manipulatives (#) 94. Sts use audio-visual equipment (#) 95. Sts use science, measuring equipment (#) 97. Sts use games and puzzles (#) 98. Sts use art materials (#)

[1]Variables are classified into higher order categories. Composite Variables represent additional ways of grouping items. Items within the Composite Variables also appear on preceding variables.

[2]For the St-DOS, some behaviors were rated according to the amount of time the behavior occurred for most of the class and some behaviors were rated according to the number of students exhibiting the behavior. Items for which the number of students were rated are indicated by (#) following the item. Note that several items are rated on both scales.

(cont.)

Table 1 (p.2)

DIMENSIONAL OCCURRENCE SCALE VARIABLES WITH COMPONENT ITEMS

Categories of Variables	Variables	Items from T-DOS	Items from St-DOS
SUPPORT STRUCTURE FOR LEARNING	ORIENTATION OF LEARNING ACTIVITY	25. T emphasizes task 20. T focuses on rote task 64. T oriented to novel 9. T emphasizes comprehension, analysis 55. T relates environment to skill and concept learning 57. T relates topics that individual Sts were interested in to their skill and concept learning 16. T allows for and/or encourages exploration	76. Superficial tasks 69. Sts work on convergent task 80. Sts work on convergent tasks (#) 75. Sts are engaged in receptive activities 71. Sts work on divergent tasks 86. Sts work on divergent tasks (#) 78. Sts tasks involve analysis, synthesis, application, evaluation 109. Sts apply learning to situations beyond immediate context
	RESPONSIBILITY FOR LEARNING	2. T allows for choices 58. T flexible regarding procedures 47. T seeks procedural suggestions 54. T accepts procedural suggestions 56. T directs class activities 15. Activities predetermined	72. Sts make choice of task(s) 104. Sts make choice of task(s) (#) 84. Sts move freely for learning purposes (#) 89. Sts show initiative (#) 90. Sts plan own sequence of activities (#) 106. Sts have voice in planning activities 70. Sts wait for and work at T/OA direction 67. Sts follow posted or stated sequence of tasks 87. Sts seek T for help (#) 91. Sts work without aid (#) 114. Sts show responsibility for materials
	EVALUATION OF LEARNING	12. T conferences with Sts to evaluate work 26. T checks individuals work as Sts are working 53. T gives direct, immediate feedback (informative) 44. T helps St analyze errors 3. T passes back corrected work 8. T checks completed work 18. T records progress	79. Sts correct each other's work 82. Sts correct own work(#) 111. Sts record own progress

(cont.)

Table 1 (p.3)

DIMENSIONAL OCCURRENCE SCALE VARIABLES WITH COMPONENT ITEMS

Categories of Variables	Variables	Items from T-DOS	Items from St-DOS
SUPPORT STRUCTURE FOR LEARNING	CONTROL	4. T uses non-verbal control 48. T redirects 61. T uses overt control 32. Noisy (chaotic) 52. T encourages internal control 41. T allows interpersonal problem solving	103. Sts move freely (#) 74. Sts talk freely 83. Unruly behavior (#)
COGNITIVE ORIENTATION	ACCEPTANCE	21. T allows for and/or encourages open expressiveness 27. T accepts St responses 37. T accepts Sts mistakes 39. T critical of Sts errors 46. T cuts off open expressiveness	
EMOTIONAL ENVIRONMENT	GENERAL TONE	1. T praising (verbal) 36. T expresses warmth non-verbally 43. T expresses warmth verbally 51. T uses humor 62. T laughs, smiles with Sts 35. Friendly 50. T impatient 42. T critical of Sts 45. T punitive 30. Rejecting	81. Sts smile (#) 115. Sts smile while working with T 105. Sts critical of others 108. Sts hostile
	AFFECT	59. T responsive to individual's affective needs 38. T elicits feelings 40. T expresses own feelings (affect) verbally	107. Sts express own feelings

(cont.)

Table 1 (p.4)

DIMENSIONAL OCCURRENCE SCALE VARIABLES WITH COMPONENT ITEMS

Categories of Variables	Variables	Items from T-DOS	Items from St-DOS
EMOTIONAL ENVIRONMENT	COOPERATION		66. Sts cooperative 96. Sts cooperative (#) 110. Sts accept a new member or members to group 112. Sts exclude new members from group 113. Sts competitive
	INVOLVEMENT		68. Sts involved in task-relevant activities 73. Sts uninvolved in class activities 88. Sts uninvolved in class activities (#) 100. Sts restless (#) 101. Sts displeased with class or activity (#)
COMPOSITE VARIABLES	ORIENTATION TOWARDS INDIVIDUAL	21. T allows for and/or encourages open expressiveness 27. T accepts St responses 37. T accepts St mistakes 7. T gives individual attention for cognitive purposes 2. T allows for choice 12. T conferences with Sts to evaluate work 26. T checks individual's work as Sts are working 59. T responsive to individual's affective needs	77. Individualized instruction 92. Individualized instruction (#) 85. Own pace (#) 72. Sts make choice of task(s) 104. Sts make choice of task(s) (#) 84. Sts move freely for learning purposes (#) 103. Sts move freely (#) (social) 89. Sts show initiative (#) 82. Sts correct own work (#) 111. Sts record own progress

(cont.)

Table 1 (p.5)

DIMENSIONAL OCCURRENCE SCALE VARIABLES WITH COMPONENT ITEMS

Categories of Variables	Variables	Items from T-DOS	Items from St-DOS
COMPOSITE VARIABLES	FLEXIBILITY	22. T uses flexible grouping for instruction 2. T allows for choices 58. T flexible regarding procedures 47. T seeks procedural suggestions 54. T accepts procedural suggestions	65. Spontaneous grouping 102. Spontaneous grouping(#) 85. Own pace (#) 72. Sts make choice of task(s) 104. Sts make choice of task(s) (#) 84. Sts move freely for learning purposes(#) 103. Sts move freely (social) (#) 89. Sts show initiative (#) 90. Sts plan own sequence of activities (#) 106. Sts have voice in planning activities

Table 2

Schedule of Observations

Week	1		2		3	
Day	1	2	1	2	1	2
Period 1	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS
Period 2	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS	T-DOS St-DOS
Period 3	T-DOS	T-DOS	T-DOS	T-DOS	T-DOS	T-DOS

Note: Periods were usually bounded by recess or lunch. Because most of the classrooms in the sample had one short 20-30 minute period during the morning, only one observation (with the T-DOS) could be made during period 3.

Table 3

Friedman's χ_r^2 and Kendall's Coefficient of Concordance (w) for Variables on the Teacher DOS for each Teacher, Within and Across Periods

		Teacher A		Teacher C		Teacher E		Teacher F		Teacher H		Teacher K	
Variable [b]	Period	χ_r^2	W[a]	χ_r^2	W	χ_r^2	W	χ_r^2	W	χ_r^2	W	χ_r^2	W
Grouping (10)	1	16.59+	.55	9.34		14.37		23.51*		15.88+	.54	35.35*	.81
	2	22.80*	.65	9.66		27.28*	.71	26.26*	.70	13.84		31.05*	.76
	3	12.21		17.21*	.57	14.03		36.22*	.82	15.35+	.28	17.49*	.57
	All	45.07*	.54	28.33*	.42	36.18*	.47	59.45*	.61	34.53*	.46	48.19*	.55
Subject Matter/ Materials(4)	1	3.35		6.45+	.60	9.35*	.72	1.90		9.90*	.74	0.00	
	2	10.95	.78	6.45+	.60	.40		9.65*	.73	1.25		14.60*	.90
	3	4.30		2.55		2.05		.40		6.95+	.39	2.60	
	All	4.25		.85		5.55		4.65		11.32*	.46	9.67*	.42
Orientation of Learning Activity (7)	1	20.20*	.75	11.34+	.56	16.02*	.67	15.16*	.65	17.41*	.70	4.25	
	2	19.95*	.74	8.30		9.30		15.02*	.65	6.39		19.36*	.73
	3	13.50*	.61	10.82+	.55	13.21*	.61	4.30		19.90**	.55	13.41*	.61
	All	52.70*	.70	29.66*	.52	35.41*	.57	31.64*	.54	47.44*	.68	30.58*	.53
Responsibility for Learning (6)	1	20.07*	.91	17.03*	.84	14.63*	.78	11.67*	.70	17.43*	.85	9.13+	.62
	2	13.43*	.75	9.07+	.62	13.13*	.74	13.43*	.75	19.10**	.64	13.73*	.76
	3	12.57*	.72	6.13		12.60*	.73	13.07*	.74	20.60**	.69	5.63	
	All	42.98*	.77	25.59*	.60	35.70*	.70	37.44*	.72	39.14*	.74	23.29*	.57
Evaluation of Learning (7)	1	24.14*	.82	3.25		26.93*	.87	13.70*	.62	17.66*	.70	3.38	
	2	13.23*	.61	8.82		.63		4.50		15.20*	.42	13.23*	.61
	3	5.91		2.09		12.25+	.58	0.00		16.54**	.46	7.02	
	All	39.32*	.60	7.81		23.76*	.47	11.14+	.32	45.38*	.65	20.48*	.44
Control (6)	1	13.81*	.68	13.10*	.66	16.07*	.73	20.95*	.84	13.81*	.68	15.00*	.71
	2	13.69*	.68	13.81*	.68	13.45*	.67	13.70*	.68	7.95	.66	15.00*	.71
	3	11.89*	.69	13.10*	.66	13.81*	.68	14.29*	.69	9.29+	.31	14.88*	.70
	All	37.12*	.66	39.05*	.66	41.23*	.68	43.14*	.69	37.61*	.67	41.15*	.68
Acceptance (Cognitive) (5)	1	18.13*	.87	6.33		12.00*	.71	6.87		9.60*	.63	15.64*	.88
	2	7.93+	.58	2.03		15.20*	.80	4.03		1.93		9.40+	.69
	3	7.40		8.33*	.59	4.80		14.47*	.78	7.83+	.33	9.13+	.62
	All	28.43*	.63	12.84*	.42	21.20*	.54	19.26*	.52	16.74*	.48	27.96*	.66
General Tone (10)	1	14.94+	.53	11.78		35.08*	.81	23.32*	.72	26.68*	.70	19.63*	.60
	2	17.47*	.57	17.31*	.57	31.63*	.77	16.66+	.56	26.22**	.42	16.80+	.56
	3	27.42*	.78	26.84*	.71	43.36*	.90	33.31*	.79	34.17**	.71	14.93+	.53
	All	55.64*	.60	48.52*	.55	104.38*	.80	67.61*	.66	79.86*	.70	44.20*	.52
Affect (3)	1	.25		.25		.25		7.00*	.76	1.75		.25	
	2	.75		0.00		1.08		0.00		.75		.75	
	3	.75		0.00		.75		.25		1.00		1.75	
	All	1.08		.08		.53		1.58		2.08		.75	
Orientation towards Individual (8)	1	27.38*	.81	15.39*	.61	23.19*	.74	14.08*	.58	21.19*	.71	21.07*	.78
	2	13.35+	.56	16.64*	.63	19.63*	.68	7.90		10.91		12.93+	.61
	3	8.58		12.39+	.54	12.13+	.54	19.35*	.68	20.61**	.42	9.06	
	All	41.33*	.57	36.58*	.54	36.81*	.54	23.09*	.43	40.22*	.57	25.72*	.48
Flexibility (5)	1	2.15		7.50+	.65	1.10		2.45		.40		0.00	
	2	2.70		7.50+	.65	2.70		2.15		2.50		.45	
	3	2.70		3.65		2.40		1.10		.33		2.70	
	All	6.32+	.34	17.22	.57	5.07		5.15		1.17		1.75	

*$p \leq .05$; +$p \leq .10$.

Note a: W is calculated only where χ_r^2 is significant.

Note b. Number of items in each variable is listed in parentheses.

Table 4

Friedman's χ_r^2 and Kendall's Coefficient of Concordance (w) for the Variables on the Student DOS for each Teacher Within and Across Periods

		Teacher A		Teacher C		Teacher E		Teacher F		Teacher H		Teacher K	
Variable[b]	Period	χ_r^2	W[a]	χ_r^2	W	χ_r^2	W	χ_r^2	W	χ_r^2	W	χ_r^2	W
Grouping (5)	1	13.70*	.76	14.63*	.78	11.70*	.70	12,93*	,73	13.33*	.75	13.73*	.76
	2	13.80*	.76	3.33		12.00*	.71	12.80*	.73	11.85*	.70	12.20*	.71
	All	25.55*	.73	11.48*	.49	19.85*	.64	25.27*	.73	24.58*	.72	23.97*	.71
Subject/ Materials (9)	1	16.81*	.35	11.44		17.76*	.37	13.07		15.69*	.33	5.86	
	2	34.60*	.72	17.38*	.36	2.91		33.47*	.70	15.90*	.33	27.67*	.57
	All	35.19*	.37	13.54+	.14	8.77		40.61*	.42	26.47*	.27	22.73*	.24
Orientation of Learning Activity (8)	1	14.89*	.84	9.65		16.82*	.63	17.85*	.65	20.53*	.70	8.60	
	2	22.99*	.74	20.56*	.70	14.92*	.65	14,18*	.64	14.60*	.59	14.50*	.64
	All	34.05*	.74	20.22*	.49	11.92		30.75*	.63	27.78*	.58	14.48*	.46
Responsi- bility for Learning (11)	1	36.06*	.85	27.65*	.68	33.17*	.74	34.17*	.76	36.07*	.78	30.91*	.72
	2	41.69*	.83	27.89*	.68	33.19*	.74	40.21*	.82	41.20*	.83	39.82*	.82
	All	74.08*	.82	35.54*	.54	58.78*	.70	71.13*	.77	74.17*	.79	66.85*	.75
Evaluation of Learning (3)	1	.00	-	.25	-	.00	-	.00	-	,25	-	.33	-
	2	2.25	-	.25	-	,00	-	.25	-	,25	-	,00	-
	All	1.13	-	.13	-	.00	-	,13	-	.50	-	,17	-
Control (2)	1	6.25+	.72	2.25	-	9,25*	,88	8.08*	,82	4,08	-	,75	-
	2	6.33+	.73	9.00*	.87	3,25	-	3.25	-	4,08	-	6,25+	.72
	All	12.54*	.72	10.13*	,65	11,63*	,70	10.79*	,67	8.17*	,58	5,38+	.47
General Tone (5)	1	9.45*	.73	7.95*	.67	14,90*	.91	9.95*	.74	14.50*	.90	6.45+	.73
	2	8.15*	.67	10.55*	.77	13.65*	.87	9.35*	.72	13.10*	.85	12.55*	.84
	All	17.28*	.69	18.13*	.71	27.63*	.88	19.05*	.73	27.30*	.87	17.79*	.77
Cooperation (5)	1	1.33	-	.00	-	2.13	-	2.03	-	1.33	-	.33	-
	2	1.33	-	10.53*	.66	.33	-	1.33	-	.50	-	3.00	-
	All	2.50	-	5.27	-	1,90	-	3.10	-	1.42	-	2.67	-
Involve- ment (6)	1	8.30*	.68	1.20	-	13.75*	,87	7.65+	.65	1.55	-	6.90+	.62
	2	7.95*	.67	4.40	-	12.15*	.82	6.05	-	13.95*	.88	8.55*	.69
	All	15.33*	.65	4.80	-	25.57*	.84	12.93*	.60	11.08*	.56	14.78*	.64
Orientation Towards In- dividual (10)	1	21.80*	.74	21.72*	.67	17.82*	.61	22.99*	.69	23.46*	.70	4.51	-
	2	25.94*	.74	24.78*	.72	13.79+	.54	18.90*	.63	26.77*	.75	24.81*	.72
	All	45.64*	.72	42.18*	.66	27.44*	.54	40.86*	.65	48.39*	.71	23.94*	.50
Flexibility (10)	1	26.59*	.77	24.15*	.67	26.36*	.70	27,16*	.71	29.02*	.73	5.57	-
	2	31.48*	.76	27.76*	.72	16.17+	.55	23.99*	,67	31.02*	.76	29.08*	.73
	All	57.36*	.76	44.26*	.64	41.25*	.62	49.22*	.68	59,06*	.74	29.04*	.52

Note a: W is calculated only where χ_r^2 is significant.

Note b: Number of items in each variable is listed in parentheses.

*$p \leq .05$

+$p \leq .10$

Table 5

Stability of Variables on the T-DOS and the St-DOS across periods as indicated by a significant χ_r^2 or Minimum Variability of Component Items

Variable	T-DOS		St-DOS	
	χ_r^2	Min.Variability	χ_r^2	Min.Variability
Grouping	*		*	
Subject/Material				
Language	*		*	
Responsibility for Learning	*		*	
Evaluation of Learning		x		x
Control	*		*	
Acceptance (T-DOS)	*		—	—
Tone	*		*	
Affect		x		
Cooperation (St-DOS)	—	—		x
Involvement (St-DOS)	—	—		x
Orientation Toward Individual	*		*	
Flexibility		x	*	

Note: An * indicates χ_r^2 of $\leq .05$.

An x indicates minimum variability of the component items.

Table 6

Stability of Variables within Periods on the T-DOS as indicated by the Significance of the χ_r^2, Kendall's Coefficient of Concordance (W), Minimum Variability of Component Items (M), and Differences in the Activity during the Period (A)

Variable	Period	Teacher A				C				E				F				H				K			
		χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^r	W	M	A
Grouping	1	+	.55						x				x	*	.66			+	.54			*	.80		
	2	*	.65						x	*	.71			*	.69							*	.75		
	3				x	*	.56		x					*	.81			+	.28			*	.56		
Subject/ Materials	1			x		+	.60		x	*	.72		x					*	.74					x	
	2	*	.78			+	.60		x			x		*	.73						x	*	.90		
	3				x				x				x				x	+	.39		x				x
Orientation of Learning Activity	1	*	.75			+	.56		x	*	.67			*	.15			*	.70					x	
	2	*	.74						x			x		*	.65					x		*	.73		
	3	*	.61			+	.55		x	*	.61						x	*	.55			*	.61		
Responsibility for Learning	1	*	.91			*	.84			*	.78			*	.70			*	.85			+	.61		x
	2	*	.75			+	.62		x	*	.74			*	.75			*	.64			*	.76		
	3	*	.72						x	*	.73			*	.74			*	.69						x
Evaluation of Learning	1	*	.82						x	*	.87			*	.62			*	.70					x	
	2	*	.61									x					x	*	.42			*	.61		
	3				x					*	.58						x	*	.46						x
Control	1	*	.69			*	.66			*	.73			*	.84			*	.68			*	.71		
	2	*	.68			*	.68			*	.67			*	.68							*	.71		
	3	*	.69			*	.66			*	.69			*	.69			+	.31			*	.70		
Acceptance (Cognitive)	1	*	.87						x	*	.71						x	*	.63			*	.88		
	2	+	.58						x	*	.80						x					+	.69	x	
	3				x	*	.59						x	*	.78			+	.33	x		+	.62		x

(cont.)

See note at end of Table.

Table 6 continued

Variable	Period	Teacher																					
		A				C				E				F				H				K			
		χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A
General Tone	1	+ .53						x		* .81				* .72				* .70				*.60			
	2	* .57				* .57				* .77				+ .57				* .42				+.56			
	3	* .78				* .70				* .90				* .79				* .71				+.53			
Affect	1			x				x				x		* .77						x				x	
	2			x				x					x			x				x			x	
	3			x				x				x				x				x			x	x
Orientation Toward Individual	1	* .81				* .61				* .74				* .58				* .71				*.78			
	2	+ .56				* .63				* .68												+.61			
	3				x	+ .54				+ .54				* .68				* .42							x
Flexibility	1			x		+ .65		x				x				x				x				x	
	2			x		+ .65		x				x				x				x				x	
	3			x								x				x				x				x	

*p < .05
+p < .10

Note: Activities contributing to variability within periods are as follows:

Teacher A: Period 3: Sharing on two days; Centers on 4 days.
Teacher C: Multiple activities occurred most of the time.
Teacher E: Period 2: One day during Sharing, affect was involved.
Period 3: Sharing on three days; Writing on 3 days.
Teacher H: Period 2: Spelling plus one other activity, e.g. Reading Writing, Art on 4 days.
Period 3: Math plus Art on one day.

Table 7

Stability of Variables Within Periods on the St-DOS as indicated by the Significance of the χ_r^2, Kendall's Coefficient of Concordance (W), Minimum Variability of Component Items (M), and Differences in the Activity during the Period (A)

Variable	Period	Teacher A				Teacher C				Teacher E				Teacher F				Teacher H				Teacher K			
		χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A	χ_r^2	W	M	A
Grouping	1	*	76			*	78			*	70			*	73			*	75			*	76		
	2	*	76							*	71			*	73			*	70			*	71		
Subject/ Materials	1	*	35					x		*	37							*	33					x	x
	2	*	72			*	36						x	*	70			*	33			*	57		
Orientation of Learning Activity	1	*	84							*	63			*	65			*	70						x
	2	*	74			*	70			*	65			*	64			*	59		x	*	64		
Responsibility for Learning	1	*	85			*	68			*	74			*	76			*	78			*	72		
	2	*	83			*	68			*	74			*	82			*	83			*	82		
Evaluation of Learning	1			x				x				x				x				x				x	
	2			x				x				x				x				x				x	
Control	1	+	72	x						*	88			*	82										x
	2	+	73			*	87									x				x		+	72		
General Tone	1	*	73			*	67			*	91			*	74			*	90			+	73		x
	2	*	76			*	77			*	87			*	72			*	85			*	84		
Cooperation	1			x				x				x				x				x				x	
	2			x		*	.66					x				x				x				x	
Involvement	1	*	68					x		*	87			+	65					x		+	62		
	2	*	67					x		*	82					x		*	88			*	69		
Orientation Toward Individual	1	*	74			*	67			*	61			*	69			*	70						
	2	*	74			*	72			+	54			*	63			*	75			*	72		
Flexibility	1	*	77			*	67			*	70			*	71			*	73						x
	2	*	76			*	72			+	55			*	67			*	76			*	73		

*p < .05, +p < .10

Note: Activities contributing to variability within periods are as follows:

Teacher E: Period 2: Sharing on 3 days, Writing on 3 days.

Teacher F: Period 1: On one day games were played.

Teacher K: Period 1: Multiple activities including art.

Table 8

Item Ratings for Selected Variables on the T-DOS for Teacher A during Period 3

Variable & Item	Activity					
	SH	SH	CE	CE	CE	CE
Grouping						
7 Indiv. Atten. Cogn.	1	1	2	3	3	3
22 Flexible GP	1	1	1	1	1	1
24 Works W. Sm. Gps.	1	1	5	1	1	3
14 Works W. Lg. Gps.	1	1	1	5	5	3
17 Pres. Info. Gps.	2	2	2	3	2	2
5 Pre-Set Gps.	1	1	5	5	5	5
11 Total Class	4	5	1	1	1	1
19 Com. Pace: Class	4	1	1	1	1	1
23 Com. Pace: Gps.	1	1	1	1	1	1
31 Rushed	1	1	0	1	1	1
Subject/Materials						
28 Varied Subjs.	1	1	5	5	5	5
29 Varied Mats.	1	1	5	5	5	5
10 Uses Text Wkbk.	1	1	1	1	1	2
6 Uses Suppl. Mat.	1	1	5	5	5	2
Evaluation of Learning						
12 Conference	1	1	1	1	1	1
26 Check As Work	1	1	1	1	2	2
53 Dir. Immed. Fdbk.	1	1	2	5	1	3
44 Anal. Errors	1	1	2	3	3	1
3 Ret. Corr. Work	1	1	1	1	1	1
8 Check Comp. Work	1	1	1	2	1	1
18 Record Progress	1	1	1	1	1	1
Control						
4 Non-Verbal	1	1	2	1	1	1
48 Redirects	1	1	1	1	3	1
61 Overt	2	2	1	1	1	1
13 Laissez-Faire	1	1	1	1	1	1
32 Noisy	3	3	0	1	1	1
52 Encour. Internal	1	1	1	1	1	1
41 Problem-Solv.	2	1	1	1	1	1
Acceptance						
21 Encour. Express	4	4	2	2	1	1
27 Accpt. St. Resp.	4	4	2	2	1	2
37 Acpt. St. Error	1	1	3	3	1	1
39 Crit. St. Error	1	1	1	1	1	1
46 Cuts Expres.	1	1	1	1	1	1

Table 9

Item Ratings for Selected Variables on the St-DOS

for Teacher E during Period 2

Variable & Items	Activity					
	SH	WR LA SP	SH	SH WR	SH	WR
Subject/Materials						
117 Use Diverse Mat.	1	1	1	1	1	1
118 Varied Subj.	1	3	1	1	1	1
116 Same Mat.	1	1	1	5	1	5
99 Use Text/Wkbk. (#)	1	1	1	1	1	1
93 Use Manips. (#)	1	1	1	1	1	1
94 Use Av. (#)	1	1	1	1	1	1
95 Use Sci./Meas. (#)	1	1	1	1	1	1
97 Use Games (#)	1	1	1	1	1	1
98 Use Art Mat. (#)	1	4	1	1	1	5
Orientation of Learning Activity						
76 Superfic. Task	1	1	1	1	1	1
69 Converg. Task	1	1	1	1	1	1
80 Converg. Task (#)	1	0	1	1	1	2
75 Receptv. Activities	5	1	5	1	5	1
71 Diverg. Tasks	1	5	5	5	5	5
86 Diverg. Tasks (#)	4	5	3	5	5	5
78 Anal. Synthesis	1	1	1	5	1	5
109 Apply Beyond	1	1	1	1	1	1
Responsibility for Learning						
72 Make Choice	1	1	1	5	1	1
104 Move Choice (#)	1	1	1	5	1	1
84 Move: Lrning. (#)	2	4	1	3	1	5
89 Initiative (#)	1	1	1	2	1	1
90 Plan Sequence (#)	1	1	1	1	1	1
106 Voice in Plan	1	2	1	2	1	1
70 Wait/Work T. Dir.	5	5	5	1	1	5
67 Posted Sequence	5	5	5	5	5	5
87 Seek T. Help (#)	1	3	1	3	1	3
91 Work Unaided (#)	1	3	1	3	1	4
114 Responsib.: Mat.	1	1	1	1	1	1

(cont.)

Table 9 (cont.)

Item Ratings for Selected Variables on the St-DOS

for Teacher E during Period 2

Variable & Items	Activity					
	SH	WR LA SP	SH	SH WR	SH	WR
Affect						
107 Express Feel.	5	1	3	1	4	1
Involvement						
68 Task Involved	5	5	5	5	4	5
73 Uninvolved	2	1	2	1	2	1
88 Uninvolved (#)	4	4	2	2	3	3
100 Restless (#)	3	3	2	1	2	2
101 Displeased (#)	1	1	1	1	1	1
Orientation Towards Individual						
77 Indiv. Instr.	1	1	1	1	1	1
92 Indiv. Instr. (#)	1	1	1	1	1	1
85 Own Pace (#)	1	5	3	5	3	5
72 Make Choice	1	1	1	5	1	1
104 Move Choice (#)	1	1	1	5	1	1
84 Move: Lrning. (#)	2	4	1	3	1	5
103 Move Freely (#)	2	4	1	3	1	5
89 Initiative (#)	1	1	1	2	1	1
82 Correct Own (#)	1	1	1	1	1	1
111 Record Progress	1	1	1	1	1	1

Table 10

Range of Ratings for Individual Items on the T-DOS for each Period for Each Teacher

Item	Period 1						Period 2						Period 3					
	Teacher						Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K	A	C	E	F	H	K
Grouping																		
7 Indiv Atten Cogn	2-5	1-3	5 4	2(5)	2-5	1	2-5	1-4	1-3	1-3	1-5	2 3	1-3	2(1,4)	1-3	1/2	2-5	1-4
22 Flexible Gp	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1 4	1	1 4	1
24 Works W Sm Gps	1(5)	1	1	1	1(4,5)	1	1-5	1	1	1	1 2	1	1(3,5)	1	1	1	1 4	1
14 Works W Lg Gps	1(5)	1-4	1	1	1	1	1-5	1	1 4	1-5	1	2-5	1-5	1 4	1 4	1	1	1 5
17 Pres Info Gps	1-3	2 1	1 2	1-5	1-4	2/3	1-3	1(4,5)	2	2-5	1-5	2/3	2 3	1-5	2(1,3)	2 5	2(1,3)	2(3,4)
5 Pre-Set Gps	5(1)	1 3	1	1 5	5(1)	1	5 1	1	1	5(1,3)	1	1-5	5(2)	1 3	1	1	1	1 5
11 Total Class	1	1-3	1	1-5	1 5	5	1	1(5)	5(2,3)	1-4	1-5	1-4	1(4,5)	1-5	2(1)	5	1(3,5)	1-5
19 Com Pace: Class	1	1-3	1	1(4,5)	1 5	5	1	1(4,5)	5(1,3)	1-4	1(4,5)	1-5	1 5	1-5	1 5	5	1-5	1-5
23 Com Pace: Gps	1 5	1(3,4)	1 5	1	1(5)	1	1 5	1	1 4	1(4)	1	1-5	1	1 4	1 4	1	1 5	1 5
31 Rushed	1	1	1	1	1	1	1	1	1	1	1 4	1	1	1 3	1	1	1 2	1
Subject/Matter																		
28 Varied Subjs	1 4	1	1(5)	1(4,5)	1 5	1	5	1/5	1 4	2-5	1-5	5(2,4)	5(1)	1-5	1-5	1 5	1 5	1(3,5)
29 Varied Mats	1 5	1	1	1 2	1	1	5	1/5	1	2-5	1 3	1	5(1)	1-5	1 2	1	1	1 3
10 Uses Text Wkbk	1-5	1-4	5	1-3	3-5	1	1 2	1	1	1 2	1(2,5)	2-5	1 2	1(3,5)	1-5	1	5(3,4)	1-5
6 Uses Suppl Mat	1 4	1(3,4)	1 2	1-5	1	1	5 4	1(3)	1 3	1-5	1 5	1	1-5	1(2,5)	1 2	1 3	1 5	1
Orientation of Learning Activity																		
25 Emph Task	5 4	2-5	5(2)	5(2)	5 4	1/5	5 4	1-5	1-5	5 2	5 4	5 2	5(1,4)	2-5	2-5	1-5	5	5(1,4)
20 Focus: Rote	1	1 3	1	1	1	1	1	1	1	1	1 5	1 2	1	1 3	1 5	1	1	1
64 Orient: Novel	1	1	1	1	1	1(2,3)	1	1	1 2	1	1	1	1	1 2	1 2	1	1	1
9 Emph Compreh	1-3	1 3	1-5	1-3	1-5	1 2	1-5	1 3	1 2	1-5	1(2,5)	1-5	1-4	1(4)	1(3,4)	4(2)	1-5	1-4
55 Environ Lrng	1	1	1	1 2	1	1	1(2,4)	1	1	1	1	1	1 2	1	1	1	1	1
57 St Interest	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1 2	1	1
16 Allow Explor	1	1 2	1 2	1	1	1 2	1 2	1 5	1 2	1	1	1	1	1 5	1	1	1	1

Note: See note at end of table.

Table 10 (cont.)

	Period 1						Period 2						Period 3					
	Teacher						Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K	A	C	E	F	H	K
Responsibility for Learning																		
2 Allow Choice	1/2	1-5	1(5)	1-5	1 2	1	1-3	5(1)	1-5	1-4	1-5	1 5	1-5	1-5	1-4	1(2)	1	1-5
58 Flex Proced	1-3	1	1(2)	1(2)	1 2	1	1	1	1(2)	1(2)	1 2	1 2	1	1-3	1 2	1(2)	1	1 2
47 Seek Proc Sug	1 2	1	1	1(2)	1	1	1	1	1	1(2)	1	1	1	1(2,3)	1	1 3	1	1
54 Acpt. Proc Sug	1(2,3)	1	1 2	1(2)	1 2	1	1	1	1(2)	1 3	1 2	1 2	1	1/2	1	1(2,3)	1	1
56 Dir Actvties	1-5	3-5	5(4)	1-5	5(2,3)	5 4	1-5	2/5	5	1-5	5(3)	5(2,4)	1-5	2-5	5(4)	3-5	5 2	5(2,4)
15 Activ Predeterm	5	5 4	5	5	5	5 1	5	1/5	5	5	5 4	5	5	5(1)	5 4	5	5	1-5
Evaluation of Learning																		
12 Conference	1	1	1	1	1	1	1	1(4)	1	1	1	1	1	1	1(2,4)	1	1	1
26 Check as Work	1/2	1	1	1-3	1/2	1	1 3	1 2	1 2	1 3	1-5	1 2	1(2)	1 2	1/2	1	1(2,4)	1(2)
53 Dir Immed Fdbk	3/5	1(4,5)	5	2-5	1-5	1-5	1-5	1-3	1 2	1-5	2-5	1-5	1-5	1 5	1-5	1	5(3,4)	1-5
44 Anal Errors	2-5	1	3-5	1(2,4)	1-5	1	1-4	1-3	1	1(2)	1-5	1-3	1-3	1(2,3)	1(2,5)	1	1-5	1 3
3 Ret Corr Work	1		1 2	1	1 2	1	1	1	1	1	1	1	1	1	1 2	1	1 2	1
8 Check Comp Work	1 3	1 2	1-3	1	2(1)	1	1 2	1/2	1	1	1/2	1 2	1 2	1	1-4	1	1	1 2
18 Record Progress	1		2(1)	1	1 2	1	1	1 2	1	1	1	1	1	1	1 2	1	1	1
Control																		
4 Non-Verbal	1	1	1/2	1 2	1	1/2	1(2)	1(2)	1 2	1(2)	1(2)	1	1 2	1	1 2	1(2)	1	1
48 Redirects	1	1 3	1-3	1	1 3	1 2	1	1 2	1(2)	1/2	1 2	1(2)	1 3	1 2	1	1 2	1(2)	1
61 Overt	1 4	1	1 3	1 2	1-4	1	1(2,4)	1(2)	1-5	1-3	1-5	1	1(2)	1 2	1 2	1-4	1-5	1 2
32 Noisy	1	1	1	1	1	1	1	1	1 2	1 2	1	1	1(3)	1	1	1	1	1 2
52 Encour Internal	1 2	1	1	1 2	1	1	1	1	1	1	1 2	1	1	1	1	1	1 2	1 2
41 Problem-Solv	1	1	1	1	1	1	1	1	1	1	1	1	1 2	1	1	1	1	1

Table 10 (cont.)

	Period 1						Period 2						Period 3					
	Teacher						Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K	A	C	E	F	H	K
Acceptance																		
21 Encour Express	1^{2}	1	1	$1^{(2)}$	1^{2}	2-5	1-4	1	1-5	1^{2}	1^{4}	1^{2}	1-4	1	1	2-5	1^{2}	1-5
27 Accpt St Resp	2-4	1/2	2-4	1-3	1-4	1-3	1-4	$1^{(2)}$	$2^{(3,4)}$	1-3	1-4	2	1-4	$2^{(1,3)}$	1-3	1-5	$2^{(1,4)}$	$2^{(3,4)}$
37 Acpt St Error	2-4	$1^{(2)}$	1	1-3	1-4	$1^{(2)}$	1-4	$1^{(2,3)}$	1	1^{3}	$1^{(3,3)}$	1/2	$1^{(3)}$	1	$1^{(3,4)}$	$1^{(2)}$	1-4	$1^{(2,3)}$
39 Crit St Error	1	1	1	1	1	1	1	1	1	1	1^{2}	1	1	1	1	1	1	1
46 Cuts Expres	1	1	1	1^{2}	1	1	1	1	1	1	1	1	1	1	1	1	1	1
General Tone																		
1 Praising	$2^{(1)}$	1^{2}	2^{3}	$2^{(1)}$	2^{1}	2	1/2	$2^{(1)}$	$2^{(1,3)}$	1/2	1-3	$2^{(1)}$	$2^{(1)}$	$1^{(1)}$	2^{3}	1/2	2	$2^{(1)}$
36 Nonverb Warm	$1^{(2,3)}$	1^{2}	1-5	1-3	1	1^{2}	$1^{(2,3)}$	1/2	1-4	1^{2}	$1^{(2)}$	$2^{(1)}$	$2^{(1,3)}$	1-3	2-5	$1^{(2)}$	1	1-5
43 Verbal Warm	$1^{(3,5)}$	1	2-5	1-3	1-5	1-3	$1^{(2,3)}$	1	1-5	$1^{(2,4)}$	1^{3}	$2^{(1)}$	$1^{(2)}$	1-3	1-5	1-5	1-3	1-4
51 Humor	1/2	1	1	1^{2}	1^{4}	$1^{(2)}$	1-3	1^{2}	1-5	1	1^{2}	1	$1^{(2)}$	$2^{(1)}$	$1^{(2,3)}$	$1^{(2)}$	1^{2}	1
62 Laugh Smile	1-5	1/2	1-5	$2^{(3)}$	$5^{(2,4)}$	$2^{(1,4)}$	1-5	1-5	1-5	1-3	1-5	$1^{(2,3)}$	2-5	1-5	3-5	3-5	2-4	1-3
35 Friendly	$5^{(1)}$	$5^{(1,3)}$	5^{4}	5^{1}	5^{1}	$1^{(5)}$	5^{1}	5^{1}	5^{3}	5^{1}	5^{1}	$5^{(1)}$	5	5^{4}	5	5	5^{2}	1/5
50 Impatient	1	1	1	1	1	1	1^{2}	1	$1^{(2)}$	1^{2}	$1^{(2)}$	1	1	1	1	1	1^{2}	1
42 Crit of Sts	1	1	1	1	1	1	1	1	1^{2}	1	1	1	1	1	1	1	1	1
45 Punitive	1^{2}	1	1^{2}	1	1^{2}	1	1^{2}	1	1	1	$1^{(2,3)}$	1	1	1	1	1	1^{2}	1
30 Rejecting	1	1	1^{2}	1	1	1	1^{2}	1	1^{2}	1^{2}	1	1	1	1	1	1	1	1
Affect																		
59 Affecv Needs	$1^{(2)}$	1^{2}	1/2	2	$1^{(2)}$	1	1^{2}	1^{2}	1^{2}	1	1^{2}	1/2	1^{2}	1^{2}	$1^{(2)}$	$1^{(2)}$	$1^{(2)}$	1-3
38 Elicit Feel	1^{2}	1	1/2	$1^{(2)}$	1	1	1	1	1^{4}	1	1	1^{2}	1	1^{2}	1^{2}	$1^{(2,5)}$	1	$1^{(2,3)}$
40 Express Feel	1^{2}	1	$1^{(2)}$	1	1/2	1^{2}	$1^{(2)}$	1	$1^{(2,5)}$	1	$1^{(2)}$	$1^{(2)}$	$1^{(2)}$	1^{2}	1/2	1/2	1	1

Note: A superscript refers to the fact that a different rating point was received on one occasion only. A superscript in parentheses refers to different rating points received on two occasions. Ratings with a slash mean that each of the rating points indicated was received on three occasions.

Table 11

Range of Ratings for Individual Items on the St - DOS for each Period for Each Teacher

Item	Period 1						Period 2					
	Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K
Grouping												
85 Own Pace (#)[b]	3-5	5	1	5	1[a]	5	1(4,5)	3-5	5	1-5	2-5	5	5[3]
77 Indiv Instr	1	1[5]	1/5	1[3]	1[4]	1		1	1(5)	1	1[5]	1[2]	1
92 Indiv Instr (#)	1-3	1(2,4)	1/5	1-3	1-4	1		1-2	1-5	1	1(2,3)	1[4]	1
65 Spont Gps	1	1	1[2]	1(3,5)	1	1		1	1	1	1	1[3]	1
102 Spont Gps (#)	1	1	1(4,5)	1(3,5)	1[2]	1		1	1-3	1-3	1(2,3)	1(2)	1[2]
Subject/Mat													
117 Use Diverse Mat	1-5	1	1-5	1-5	1-5	1		5	5[1]	1	5	1-5	1[2]
118 Varied Subj	1	1[3]	1(5)	1-5	1(5)	1[4]		5	5(2)	1[3]	5	1-5	5
116 Same Mat	1[2]	1-5	1	1[3]	1-5	1-5		1	1[5]	1(5)	1	1-4	1
99 Use Text/Wkbk (#)	5[3]	1-5	5[4]	1-5	5[3]	1-5		4(2,3)	1-4	1	2-5	1-5	3-5
93 Use Manips (#)	1(2)	1	1	1-3	1[2]	1		1-4	1	1	1(2,4)	1(2)	1
94 Use Av (#)	1	1	1	1	1	1		1[2]	1	1	1	1	1
95 Use Sci/Meas (#)	1	1	1	1[2]	1	1		1	1/3	1	1(2)	1	1[2]
97 Use Games (#)	1(2,3)	1	1	2(1,5)	1[3]	1		1	1[2]	1	1-3	1	1
98 Use Art Mat (#)	1	1	1[2]	1(2,5)	1[2]	1[2]		1-3	1-5	1(4,5)	1-4	1-5	1-3

Note: See note at end of table.

(cont.)

Table 11 (cont.)

Range of Ratings for Individual Items on the St-DOS for each Period for Each Teacher

Item	Period 1						Period 2					
	Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K
Orientation of Learning Activity												
76 Superfic Task	1	1	1	1 2	1	1-3	1	1	1	1	1	1
69 Converg Task	5	1-5	5 1	1-5	5 3	1-5	5 3	1 (4,5)	1	5 (1)	1-5	5 1
80 Converg Task (#)	5 4	5 (1)	5 1	2-5	5 4	1-5	4 (3,5)	1-5	1 2	5 (2,3)	1-5	1-5
75 Receptv Actvties	1 (3,4)	1-4	1	1	1 (3,4)	1-5	1	1	1/5	1	1-3	1 2
71 Diverg Tasks	1	1 (2,4)	1 3	1-5	1 3	1 4	1 4	5 (3,4)	5 1	1 (5)	1-5	1 (4,5)
86 Diverg Tasks (#)	1 3	1 (5)	1 (2,4)	1-5	1 (2,4)	1 4	1-5	3-5	5 (3,4)	2-4	1-5	1-5
78 Anal, Synthesis	5 (3,4)	1-5	1-5	1-5	1 (4,5)	1 (2,4)	1/5	1/5	1 (5)	1 (5)	1 (3,4)	5 (1)
109 Apply Beyond	1	1	1	1	1	1 2	1	1	1	1	1	1 3
Responsibility for Learning												
72 Make Choice	1 2	1	1 5	1 3	1	1 4	1	5 (1)	1 5	1	1 (2,3)	1 4
104 Move Choice (#)	2 (1)	1	1 (2,5)	1-3	2 (1)	1	1 2	5 (1)	1 5	1 (3)	1-5	1 2
84 Move: Lrning (#)	2-5	1-5	1-5	2-5	1-5	1-5	2-5	5 (2,4)	1-5	1-5	2-5	2-5
89 Initiative (#)	1	1	1	1 3	1 2	1	1	1	1 2	1	1/2	1 2
90 Plan Sequence (#)	1 2	1	1	1 2	1	1	1	1-5	1	1	1 2	1 3
106 Voice In Plan	1 2	1	1	1 2	1/2	1	1	1 3	1 (2)	1	1 (2,3)	1

(cont.)

Table 11 (cont.)

Range of Ratings for Individual Items on the St - DOS for each Period for Each Teacher

Item	Period 1						Period 2					
	Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K
Responsibility for Learning (cont.)												
70 Wait/Work T Dir	5	5^{1}	1-5	1-5	5	5	5	1/5	$5^{(1)}$	5	5	5^{2}
67 Posted Sequence	5	5	5	5^{4}	5	5^{4}	5	$1^{(5)}$	5	5	5^{4}	5
87 Seek T Help (#)	1-4	1-5	$3^{(2)}$	1-4	2-5	$1^{(2)}$	1-3	2-4	1/3	2	$4^{(2,3)}$	2
91 Work Unaided (#)	$1^{(2)}$	1-4	$2^{(3)}$	2-4	$1^{(4,5)}$	1-5	1-3	2-5	1-4	1-4	$2^{(5)}$	$3^{(2)}$
114 Responsib: Mat	1	1	1	$1^{(2,3)}$	$1^{(2,3)}$	1	$1^{(2)}$	1^{2}	1	1^{4}	1	1
Evaluation												
79 Correct Others	1	1	1	1	1	1^{2}	1	1	1	1	1	1
82 Correct Own (#)	1	1^{4}	1	1	1^{2}	1^{5}	1/2	1	1	1	1^{2}	1
111 Record Progress	1	1	1	1	1	1	1	1^{2}	1	1^{3}	1	1
Control												
103 Move Freely (#)	1-5	1-5	2-5	2-5	1-5	$1^{(2,5)}$	1-3	2-5	1-5	1-5	1-5	1-5
74 Sts Talk Freely	5	1-5	2-5	5	1/5	1-4	5	5	1-5	5	2-4	1-5
83 Unruly Behavior(#)	1	1	1^{2}	1	1	1	1	1	1	1	1/2	1

(cont.)

Table 11 (cont.)

Range of Ratings for Individual Items on the St - DOS for each Period for Each Teacher

Item	Period 1						Period 2					
	Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K
General Tone												
81 Smile (#)	1-4	2/3	3-5	1-5	2-4	$2^{(3,5)}$	2-5	2-4	2-5	$2^{(3)}$	2-5	2^{3}
115 Smile: W T	1-5	1-3	2-5	$2^{(3,5)}$	2-5	$1^{(2,5)}$	1-5	1-3	2-5	1-3	2-5	2^{1}
105 Crit Others	1	1^{3}	1	$1^{(2)}$	1^{2}	1	$1^{(2,3)}$	1	1^{2}	1^{3}	1-3	1
108 Hostile	1^{2}	1^{2}	1^{2}	1	1	1	$1^{(2)}$	1	1	1	1	1
Affect												
107 Express Feel	1/2	1^{3}	1^{2}	1^{3}	1^{2}	$1^{(2,5)}$	1-3	$1^{(2,3)}$	1-5	1^{3}	$1^{(2)}$	1
Cooperation												
66 Cooperative	1	1	1	1	1	1	1	1^{5}	1	1	1	1
96 Cooperative (#)	1	1	$1^{(3)}$	1/2	$1^{(2)}$	1^{2}	1^{2}	2-4	1^{2}	$1^{(2)}$	1^{2}	1/2
110 Accpt Mems	1	1	$1^{(2)}$	1^{2}	1	1	1	1	1	1	1^{3}	1
112 Exclude Members	1	1	1	1^{2}	1	1	1	1	1	1^{2}	1	1
113 Competative	$1^{(2)}$	1	1	1^{5}	1	1	$1^{(2)}$	1	1	1	1	1
Involvement												
68 Task Involved	5	5	2-5	$5^{(4)}$	$5^{(1,4)}$	$5^{(3,4)}$	$5^{(4)}$	5	5^{4}	5^{4}	$4^{(5)}$	5

(cont.)

Table 11 (cont.)

Range of Ratings for Individual Items on the St - DOS for each Period for Each Teacher

Item	Period 1						Period 2					
	Teacher						Teacher					
	A	C	E	F	H	K	A	C	E	F	H	K
Involvement (cont.)												
73 Uninvolved	1	1	1-4	$1^{(2,3)}$	1^{2}	1^{5}	1^{2}	1	1/2	1^{2}	$2^{(1)}$	1
88 Uninvolved (#)	2	$1^{(2,3)}$	3-5	1-4	1-3	$2^{(1,4)}$	1-4	$2^{(1)}$	2-4	1-3	2-4	$2^{(1,4)}$
100 Restless (#)	1^{2}	1	2-4	1/2	$1^{(2)}$	1-4	1-3	$1^{(2)}$	1-3	$2^{(1)}$	2-4	1-3
101 Displeased (#)	1^{2}	1	1	1	$1^{(2)}$	1	1	1	1	1	1^{2}	1

Note a: Superscript refers to the fact that a different rating point was received on one occasion only. A superscript in parentheses refers to different rating points received on two occasions. Ratings with a slash mean that each of the rating points indicated was received on three occasions.

Note b: # indicates the item was rated on Number of Students Rating Scale.

APPENDIX

DOS: Interrater Agreement*

Interrater agreement was calculated using Ebel's intraclass correlation (Guilford, 1954). After the data was gathered, the three observers simultaneously rated one live class situation and three videotapes of three different teachers.

It should be noted that estimating interrater agreement by this method has two types of limitations: (1) limitations due to the statistical technique used, and (2) limitations due to the number and type of situations on which the calculations were based.

Limitations due to Statistical Technique

As noted by Frick and Semmel (1974), the intraclass correlation is highly affected by the variance between teachers. To obtain a correlation coefficient of 1.00, all observers must agree on the rating of each situation. With little variation between teachers, mild differences among the raters will be accentuated. Conversely, where there are large differences between teachers on a given type of behavior, differences between raters will be minimized. In addition, if teachers are rated similarly with only one or two differences among the raters, the intraclass correlation coefficient may be negative. Consequently, the meaning of the intraclass correlation coefficient is not always consistent with the usual meaning of interrater agreement.

Limitations due to Number and Type of Situations

Because only four teachers could be observed for purposes of calculating interrater reliability, and only one of these was in an actual ongoing classroom situation, the estimate of reliability based on this data should be taken with caution. The small number in the sample does not fully meet the assumptions

*The issue of interrater reliability is considered in the Interim Report (Marshall, 1973).

necessary for the analysis of variance. In addition, the videotapes used were restricted to the extent that (1) they centered on the teacher rather than the whole class and (2) this structure differed from those observed during the actual data collection. However, because these teachers were not very diverse in traits or in classroom structure, the coefficient of interrater agreement probably overestimates observer disagreement.

Additional Methods of Estimating Interrater Agreement

Because of these limitations, a measure of explained variance due to rater differences was calculated (Marascuilo, 1971). In addition, the raw data was inspected and the number of times the observers varied by one, two, three, and four points was recorded for each item. Table A displays (1) the intraclass correlation coefficients, (2) the measure of explained variance (omega squared), and (3) the number of times the raters disagreed for each item.

Interrater Agreement

Of the 64 items on the T-DOS for which a correlation was calculated, five items (#4, #27, #33, #34, #43, #53) were associated with an explained variance (ω^2) due to rater difference equal to or greater than 30%. Two of the 54 items (#101, 105) on the St-DOS had an explained variance due to rater differences of greater than 30%. On first glance, these items would appear suspect. However, on two of the T-DOS items (#4, #43), the raters were off by only one point twice. On item #27, the raters disagreed by one point five times. On the two St-DOS items, the raters were off by only one point three times. Consequently, the observer differences on these items need not be considered as divergent as the explained variance would imply. On item #27, raters disagreed by one point five times: on item #53, the rater disagreed by one point eleven times; and three points once. These items were retained in the data analysis.

Items Eliminated from Data Analysis

As a further check on the reliability of the items, the observers were asked which items they had the most difficulty coding. These items were: #3 Efficient, #34 Calm, #60 Patient, #63 Tolerant. Based on the experience of the observers, the decision was made to eliminate these items from the actual data analysis. Two of the four items (#33 and #34) were two items for which the explained variance exceeced 30% and for which there was more than minimal observer disagreement.

Two additional items were excluded from the data analysis, since these behaviors were never observed during the data collection: #13 Laissez faire, and # 49 Cuts off exploration.

Summary of Interrater Agreement

After the elimination of the six items from the T-DOS noted above, the three raters agreed perfectly on 33 of the remaining 58 items (57% of the items), although on 18 items no occurrence of the behavior was observed.* On eight items (14% of the items), the observers disagreed once by one point. On 10 items (17%), the observers disagreed twice by one point. Totalling these percentages indicates that on 88% of the items, the observers disagreed by one point two or less times--out of a possible twelve.** On 5 items (9%), the observers disagreed once by two points or by one or two points once. Again, totalling these percentages indicates that on 96% of the items, the observers disagreed two or less times--out of a possible twelve-by two or less points.

For the St-DOS, the observers agreed perfectly on 30 of the 54 items (56%)--although on 19 items no occurrence of the behavior was observed. On seven items (13%), they disagreed once by one point. On four items (7%), they disagreed by one point twice. Summing these percentages reveals that on 79% of the items, the

* Note that the observers agreed that these behaviors did not occur.

** The twelve possible points of disagreement result from three raters observing four teachers.

observers disagreed less than three times--out of a possible 12--by two or less points.

TABLE A

INTRACLASS CORRELATIONS, EXPLAINED VARIANCE, AND NUMBER OF TIMES OBSERVERS DIFFERED ON EACH ITEM ON THE DOS

ITEM	r_{KK}	ω^2	Number of Points by which Observers Differed 1	2	3	4
T-DOS						
1	.02	.18	1[3]			
2	.87	.04	1	1		
3	1.00[1]					
4	- .02	.40	2			
5	1.00[1]					
6	.96	.02		1		
7	.93	.05	3			
8	.54	.22	2			
9	.27	.08	2			
10	1.00					
11	1.00					
12	1.00[1]					
13	1.00[1]					
14	.94	.04				
15	1.00					
16	1.00[1]					
17	.94	.13	3			
18	1.00[1]					
19	1.00					
20	0.00	.18		1		
21	.75	.10	1			
22	1.00[1]					

See note at end of Table.

(cont.)

ITEM	r_{KK}	ω^2	Number of Points by which Observers Differed 1	2	3	4
T-DOS (con't)						
23	1.00[1]					
24	.75	.10	1			
25	1.00					
26	.66	.12	2			
27	.75	.50	5			
28	1.00[1]					
29	1.00[1]					
30	1.00[1]					
31	1.00[1]					
32	1.00[1]					
33	.83	.30	2	2	1	
34	.87	.41	3	2		
35	.98		2			
36	.84	.03	2			
37	.55	.22	2			
38	1.00[1]					
39	1.00[1]					
40	1.00					
41	1.00[1]					
42	.89	.06	1	1		
43	- .01	.40	2			
44	.67	.12	1			
45	.94	.04	1			
46	1.00[1]					
47	.75	.10	1			
48	.67	.12	2			

(cont.)

ITEM	r_{KK}	ω^2	Number of Points by which Observers Differed 1	2	3	4
T-DOS (con't)						
49	1.00[1]					
50	.96	.02	2			
51	.60	.10	4			
52	- .01	.18	1			
53	- .76	.53	3		1	
54	1.00					
55	1.00[1]					
56	0.00	.18		1		
57	1.00[1]					
58	1.00					
59	1.00					
60	1.00[1]					
61	.99	.02	2			
62	.98	.05	3			
63	1.00[1]					
64	1.00[1]					
S-DOS						
65	1.00[1]					
66	1.00[1]					
67	1.00					
68	.80	.15	2		1	
69	1.00					
70	1.00					
71	1.00					
72	1.00					

(cont.)

ITEM	r_{KK}	ω^2	Number of Points by which Observers Differed 1	2	3	4
S-DOS (con't)						
73	.80	.13	2		1	
74	.79	.06	2			1
75	.59	.15	3	1		
76	1.00[1]					
77	1.00[1]					
78	- .87	.10			1	1
79	1.00					
80	1.00					
81	.95	.14	2	1	1	
82	1.00[1]					
83	0.00	.18	1			
84	.94	.04	1			
85	1.00					
86	1.00					
87	.92	.08	2			
88	.93	.04	3			
89	1.00[1]					
90	1.00[1]					
91	.92	.05	2			
92	1.00[1]					
93	0.00	.18				1
94	1.00[1]					
95	1.00[1]					
96	0.00	.18	1			
97	1.00[1]					
98	0.00	.18				1

ITEM	r_{KK}	ω^2	Number of Points by which Observers Differed 1	2	3	4
S-DOS(con't)						
99	.99	.00	1			
100	.93	.15	2	2		
101	.36	.44	3			
102	1.00[1]					
103	.96	.02	1			
104	1.00[1]					
105	.54	.40	3			
106	1.00					
107	.75	.10	1			
108	- .01	.18	1			
109	1.00[1]					
110	1.00[1]					
111	1.00[1]					
112	1.00[1]					
113	.72	.11	1	1		
114	1.00[1]					
115	.99	.01	2			
116	.95	.03	1	1		
117	1.00[1]					
118	1.00[1]					

Note 1. No occurrences of this behavior were observed by any of the raters on any teacher.

Note 2. Numbers in the columns refer to the number of times the raters disagreed by 1, 2, 3, 4 points.

Note 3. Numerals in these columns indicate the number of times observers disagreed by one, two, three, and four rating points.